

Abstract

A method and system are disclosed for organizing and retrieving information through the use of taxonomies, a document classifier, and an autocontextualization system.

5 According to the present invention, documents stored in the organization and retrieval subsystem may be manually or automatically classified into a predetermined number of taxonomies. In operation the documents are first transformed from clear text into a structured record and to automatically construct indexes to help identify when the structured record is an appropriate response to a particular query. An automatic term extractor creates a list of terms that are indicative of the subject matter contained in the documents, and then a subject matter expert identifies the terms that are relevant to the taxonomies. A term analysis system assigns the relevant terms to one or more taxonomies, and a suitable algorithm is then used to determine the relatedness between each list of terms and its associated taxonomy. The system then clusters documents for each taxonomy in accordance with the weights ascribed to the terms in the taxonomy's list and a directed acyclic graph (DAG) structure is created.

10
15
20 The present invention may then be used to aid a researcher or user in quickly identifying relevant documents, in response to an inputted query. It may be appreciated that both a document's content and information added during autocontextualization is available for retrieval in the present invention. Moreover, the present system can retrieve any type of knowledge container, including not only those derived from some kind of document (such as "document" or "question" knowledge containers) but also those that

le and resources (such as kno

W **E** **N** **I** **S**

Serial No. 09/594,083

SYSTEM AND METHOD FOR IMPLEMENTING A KNOWLEDGE MANAGEMENT SYSTEM

Abstract

A method and system organize and retrieve information using taxonomies, a document classifier, and an autocontextualizer. Documents (or other knowledge containers) in an organization and retrieval subsystem may be manually or automatically classified into taxonomies. Documents are transformed from clear text
5 into a structured record. Automatically constructed indexes help identify when the structured record is an appropriate response to a query. An automatic term extractor creates a list of terms indicative of the documents' subject matter. A subject matter expert identifies the terms relevant to the taxonomies. A term analysis system assigns the relevant terms to one or more taxonomies, and a suitable algorithm is then used to
10 determine the relatedness between each list of terms and its associated taxonomy. The system then clusters documents for each taxonomy in accordance with the weights ascribed to the terms in the taxonomy's list and a directed acyclic graph (DAG) structure is created.